Brenier Maps to Knothe-Rosenblatt Map

Medha Agarwal & Garrett Mulcahy

October 11 2022

Preliminaries

The problem of characterizing complex measures is often dealt by transporting them to a simpler target measure via a transport map T. Let $\mu : \mathcal{B}(\mathbb{R}^d) \to \mathbb{R}^+$ be the source measure defined on the Borel σ-algebra of $\mathbb{R}^d$ and let $\nu : \mathcal{B}(\mathbb{R}^d) \to \mathbb{R}^+$ be the target measure we wish to transport μ to. Then the transport map T pushes μ to ν if $T_{\#}\mu = \nu$, i.e. for any $A \in \mathcal{B}(\mathbb{R}^d)$,

$$T_{\#}\mu(A) := \mu(T^{-1}(A)) = \nu(A). \tag{1}$$

The transport map is chosen based on the cost it takes to transport μ to ν. It is characterized by the cost function $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$ which leads to Monge problem

$$\inf_{T:T_{\#}\mu=\nu} \int_{\mathbb{R}^d} c(x, T(x)) d\mu(x) \,. \tag{2}$$

The solution to the constrained problem in (2) is the optimal transport from μ to ν.

Brenier maps. Brenier maps arise from a relaxation of (2) called Monge-Kantorovich problem and defined as

$$\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) d\pi(x, y), \tag{3}$$

where $\Pi(\mu, \nu)$ is the subspace of $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ (the set of Borel probability measures on $\mathbb{R}^d \times \mathbb{R}^d$) with marginals equal to μ and ν, respectively. Theory regarding the solutions of (3) exists in greater generality than stated, but for the time being we will just state the results that we will need and give appropriate citation.

First, if $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ is lower semicontinuous in (3), then there exists $\gamma^* \in \Pi(\mu, \nu)$ such that the infimum is attained (Villani (2003), Theorem 1.3). In the case that $c(x, y) = |x - y|^2$ is the quadratic cost, with some additional assumptions we get Brenier's Theorem:

Theorem 1. *Brenier's Theorem (Villani (2003), Theorem 2.12) Let μ, ν be probability measures on $\mathbb{R}^d$ with finite second moments, and assume that μ does not give mass to small sets (this is satisfied, for example, if $\mu << m$). Then there is a unique optimal solution π^* to (3) with quadratic cost of the form $\pi^* = (\mathrm{Id} \times \nabla\varphi)\#\mu$ where $\nabla\varphi$ is the unique (up to μ-a.e. equivalence) gradient of a convex function that pushes forward μ to ν, i.e. $\nu = \nabla\varphi_{\#}\mu$.*

We call the transport map $\nabla\varphi$ arising in Brenier's Theorem the *Brenier map.* We note the degenerate support of Brenier maps compared to more general maps in $\Pi(\mu, \nu)$. Additionally, Brenier's Theorem demonstrates that, in the case of quadratic costs, solutions to the Monge-Kantorovich problem are solutions to the classical Monge problem.

An important consequence of Brenier's Theorem is that it allows us to quickly uncover solutions to the 1-dimensional optimal transport problem as by the uniqueness of $\nabla\varphi$ it suffices to find a transport map that is the gradient of a convex function. To wit, we have

Theorem 2. *Optimal Transport on $\mathbb{R}$ Let μ and ν be probability measures on $\mathbb{R}$ with finite second moments and distribution functions F and G, respectively. Additionally, assume μ is absolutely continuous. Then $G^{-1} \circ F$ is the Brenier map transporting μ to ν, where G^{-1} is the generalized inverse of G on $[0,1]$, i.e. $G^{-1}(a) = \inf\{x \in \mathbb{R} : G(x) > a\}$.*

First, it is clear that F and G^{-1} are increasing, so $G^{-1} \circ F$ is an increasing function. The function $H(x) = \int_0^x (G^{-1} \circ F)(t)dt$ is then a convex function. Thus, if we can show that $G^{-1} \circ F$ is a transport map, then by the uniqueness result in Brenier's theorem we have that $G^{-1} \circ F$ is indeed the Brenier map. Let $X \sim \mu$, then it is well-known that $F(X) \sim \text{Unif}[0,1]$. For all $t \in \mathbb{R}$ we then have

$$P((G^{-1} \circ F)(X) \leq t) = P(F(X) \leq G(t)) = G(t).$$

Thus, since $(G^{-1} \circ F)(X)$ has the same distribution as G, we have that $G^{-1} \circ F$ is indeed a transport map from μ to ν. As argued earlier, this implies that $G^{-1} \circ F$ is the Brenier map.

Knothe-Rosenblatt maps. The quadratic cost discussed above gives a unique but dense Brenier map. The denseness of Brenier maps makes operations like inversion computationally prohibitive, especially in high-dimensional settings. An alternative coupling, separately proposed by Knothe (1957) and Rosenblatt (1952), is the Knothe-Rosenblatt rearrangement which is defined when μ is absolutely continuous. The exact construction of KR maps is described by Villani (2009) and omitted here for brevity. A key quality of KR maps is the it is triangular in that its Jacobian is a triangular matrix. This makes KR maps amenable to various geometric applications. KR map can be intuitively understood as one-dimensional monotonic transformation of the marginal distribution of last coordinate, and then the sequential conditional distributions. However, such a triangular map imposes uniqueness in terms of the coordinate reordering, in contrast to optimal transport which is invariant to isometries on $\mathbb{R}^d$. In fact, if μ and ν are absolutely continuous measures, then KR map is the unique triangular map satisfying (1).

Carlier et al. (2010) show that the optimal solutions to the Monge-Kantorovich problems with weighted quadratic costs converge to Knothe-Rosenblatt maps if the weights dominate one another.

Main Result

We will assume that the Knothe-Rosenblatt map is well defined. This is true if the source measure is absolutely continuous. The precise assumption is the following.

Assumption 1. *(H-source) The marginal one-dimensional measure μ^d (dth marginal measure of μ) has no atoms. Similarly for $k \geq 1$, the one-dimensional conditional measures $\mu^k_{x_{d:k+1}}$ for $\mu^d -$ a.e. x_d, $\mu^{d-1}_{x_d} -$ a.e. $x_{d-1}, \dots \mu^{k+1}_{x_{d:k+1}}$ have no atoms.*

As opposed to Brenier's Theorem, we will also require more stringent assumptions on the target measure. This is because it will allow us to assume certain maps are invertible in the main proof.

Assumption 2. *(H-target) The marginal one-dimensional measure ν^d has no atoms. Similarly for $k \geq 2$, the one-dimensional conditional measures $\nu^k_{y_{d:k+1}}$ for ν^d- a.e. $y_d, \nu^{d-1}_{y_d}-$ a.e. $y_{d-1}, \ldots \nu^{k+1}_{x_{d:k+1}}$ have no atoms.*

We observe that (H-target) is a slightly less restrictive assumption than (H-source) as it does require any conditional of the form $\nu^1_{y_{d:2}}$ to have no atoms.

We now state the main theorem of interest.

Theorem 3. *(Carlier et al. (2010), Theorem 2.1) Let μ and ν be two probability measures on $\mathbb{R}^d$ satisfying (H-source) and (H-target), respectively, with finite second moments, and γ_ϵ be an optimal transport plan for the cost $c_\epsilon(x, y) = \sum_{i=1}^d \lambda_i(\epsilon)(x_i - y_i)^2$, for some weights $\lambda_k(\epsilon) > 0$. Suppose that for all $k \in \{1, \ldots, d-1\}$, $\lambda_k(\epsilon)/\lambda_{k+1}(\epsilon) \to 0$ as $\epsilon \to 0$. Let T_K be the Knothe-Rosenblatt transport from μ to ν and $\gamma_K \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ be the associated transport plan (i.e. $\gamma_K := (id \times T_K)_\# \mu$). Then $\gamma_\epsilon \rightharpoonup \gamma_K$ as $\epsilon \to 0$.*

Moreover, should the plans γ_ϵ be induced by transport maps T_ϵ, then these maps would converge to T_K in $L^2(\mu)$ as $\epsilon \to 0$.

The result is significant as it suggests a limiting procedure in which solutions to (3) converge weakly to the KR map. It is important to note that the sense in which the γ_ϵ are optimal is different from the sense in which the KR map is optimal, yet they are connected via this result. While γ_ϵ are solutions to Monge-Kantorovich problem expressed through certain cost functions, KR maps are not derived to minimize a pre-defined cost. However, this result ties KR maps to a cost function in an asymptotic sense.

Once a weak limit to the γ_ϵ is given (denoted γ), the main strategy of the proof is to use the optimality and uniqueness of the 1-dimensional transport maps on which the KR map (denoted γ_K) is built to iteratively show that $\gamma_K^d = \gamma^d$, then $\gamma_K^{d-1} = \gamma^{d-1}$ (the marginal on $(x_{d-1}, x_d), (y_{d-1}, y_d)$), and so on until $\gamma_K = \gamma$.

Proof of Main Result

We start with the following Lemma that will prove essential in the main argument argument, although this result is stated in greater generality than we will need.

Lemma 1. *Let μ, ν be Borel probability measures on $\mathbb{R}^d$. If $\int \phi(x) d\mu(x) = \int \phi(x) d\nu(x)$ for all $\phi \in C_c(\mathbb{R}^d)$, then $\mu = \nu$ as measures.*

Proof. The Riesz Represtation Theorem gives a 1-1 correspondence between Radon measures on a space and the dual space of $C_c(\mathbb{R}^n)$ given by $\mu \mapsto \int f d\mu$ for $f \in C_c(\mathbb{R}^n)$. Since $\mathbb{R}^n$ is a Polish space (separable Banach space), probability measures on $\mathbb{R}^n$ are Radon measures. Thus, as the actions of μ and ν on $C_c(\mathbb{R}^n)$ are the same, we must have that $\mu = \nu$, as desired. □

Lemma 2. *If two measurable functions f, g on $\mathbb{R}^d$ are equal π-almost everywhere, then $\int [f(x) - g(x)] d\pi(x) = 0$.*

Proof. The two function f and g are equal almost everywhere implies that the probability of the set on which they are unequal has a measure zero under π. Let $A \in \mathcal{B}(\mathbb{R}^d)$ be the set such that for all $x \in A, f(x) \neq g(x)$. This implies that $\pi(A) = 0$ and

$$\begin{aligned} &\int_A (f(x) - g(x)) d\pi(x) = 0 \\ \implies &\int_{\mathbb{R}^d} (f(x) - g(x)) d\pi(x) = 0 \end{aligned}$$

□

We now begin with our proof of the main result.

Proof. We will split the proof into two separate claims that will help us construct the proof of $\gamma = \gamma_K$ by equating its marginals iteratively. The first objective is to prove that the dth marginal of γ (denoted by γ^d) is equal to that of γ_K (denoted by γ_K^d). Secondly we will prove that the conditional measure of $(d-1)$th coordinate conditioned on (x_d, y_d) is also equal for γ and γ_K, i.e. $\gamma_{(x_d,y_d)}^{(d-1)} = \gamma_{(x_d,y_d),K}^{(d-1)}$.

Claim: $\gamma^d = \gamma_K^d$.

First, by dividing each c_ϵ by $\lambda_d(\epsilon)$, we can insist that $\lambda_d(\epsilon) = 1$ for all $\epsilon > 0$. For each $\epsilon > 0$, Theorem 1.3 in Villani (2003) gives the existence of an optimal $\gamma_\epsilon \in \Pi(\mu, \nu)$ for the Monge-Kantorovich problem with cost function c_ϵ. Since $\Pi(\mu, \nu)$ is sequentially compact in the topology of weak convergence, up to considering a subsequence there is a $\gamma \in \Pi(\mu, \nu)$ such that $\gamma_\epsilon \rightharpoonup \gamma$. The goal now is to show that $\gamma = \gamma_K$.

Since γ_K is a map that transports μ to ν, for all $\epsilon > 0$ the optimality of γ_ϵ with respect to c_ϵ gives that

$$\int c_\epsilon d\gamma_\epsilon \leq \int c_\epsilon d\gamma_K \,. \tag{4}$$

As $\lambda_k(\epsilon)/\lambda_{k+1}(\epsilon) \to 0$ as $\epsilon \to 0$, this implies that $\lambda_k(\epsilon) \to 0$ for all $k \in \{1, \ldots, d-1\}$. Also, since both measure μ and ν are assumed to have bounded second moments, all measures $\pi \in \Pi(\mu, \nu)$ have bounded second moment. Let $\int |x-y|^2 d\pi(x,y) \leq M$ for a constant M. Notice that for any $i \in \{1, \ldots, d-1\}$,

$$\lambda_i(\epsilon) \int |x_i - y_i|^2 d\gamma_\epsilon(x,y) \leq \lambda_i(\epsilon) \int |x-y|^2 d\gamma_\epsilon(x,y) \leq \lambda_i(\epsilon) M \to 0$$

as $\epsilon \to 0$. Let $c^{(d)}(x,y) = |x_d - y_d|^2$, then it follows by the definition of weak convergence (along with a limiting argument) that $\int c^{(d)} d\gamma_\epsilon \to \int c^{(d)} d\gamma$. This then gives

$$\int |x_d - y_d|^2 \, d\gamma \leq \int |x_d - y_d|^2 \, d\gamma_K.$$

Since each integral is just a function of (x_d, y_d), letting $\pi_d(x,y) = (x_d, y_d)$ this implies that

$$\int |x_d - y_d|^2 \, d(\pi_d)_\# \gamma \leq \int |x_d - y_d|^2 \, d(\pi_d)_\# \gamma_K. \tag{5}$$

However, since $\gamma, \gamma_K \in \Pi(\mu, \nu)$ we have that $(\pi_d)_\# \gamma$ and $(\pi_d)_\# \gamma_K$ are both transport maps from μ_d to ν_d. However, the construction of the KR map guarantees that $(\pi_d)_\# \gamma_K$ is the optimal transport map. By the uniqueness of Brenier maps, (5) forces that $(\pi_d)_\# \gamma = (\pi_d)_\# \gamma_K$. We denote this common measure γ^d.

Claim: $\gamma^{(d-1)}_{(x_d,y_d)} = \gamma^{(d-1)}_{(x_d,y_d),K}$ **for** γ^d**-a.e.** (x_d, y_d)

We now build off of this equality on the dth marginal to demonstrate equality of the joint distributions on the last two coordinates of γ and γ_K. As by construction $(\pi_d)_\#\gamma_K$ is the optimal transport map of μ_d to ν_d with respect to quadratic cost, it follows for all $\epsilon > 0$ that

$$\begin{aligned}\int |x_d - y_d|^2 \, d(\pi_d)_\#\gamma_K(x_d, y_d) + \sum_{i=1}^{d-1} \lambda_i(\epsilon) \int (x_i - y_i)^2 d\gamma_\epsilon & \\ \le \int |x_d - y_d|^2 \, d(\pi_d)_\#\gamma_\epsilon(x_d, y_d) + \sum_{i=1}^{d-1} \lambda_i(\epsilon) \int (x_i - y_i)^2 d\gamma_\epsilon & \\ = \int c_\epsilon d\gamma_\epsilon & \\ \le \int c_\epsilon d\gamma_K & \\ = \int |x_d - y_d|^2 \, d(\pi_d)_\#\gamma_K(x_d, y_d) + \sum_{i=1}^{d-1} \lambda_i(\epsilon) \int (x_i - y_i)^2 d\gamma_K. &\end{aligned}$$

Getting rid of the $\int |x_d - y_d|^2 \, d(\pi_d)_\#\gamma_K(x_d, y_d)$ term from the first and last lines then gives

$$\sum_{i=1}^{d-1} \lambda_i(\epsilon) \int (x_i - y_i)^2 d\gamma_\epsilon \le \sum_{i=1}^{d-1} \lambda_i(\epsilon) \int (x_i - y_i)^2 d\gamma_K.$$

Dividing by $\lambda_{d-1}(\epsilon)$ and letting $\epsilon \to 0$, by the same limiting argument of the previous paragraph we get that

$$\int |x_{d-1} - y_{d-1}|^2 \, d\gamma \le \int |x_{d-1} - y_{d-1}|^2 \, d\gamma_K.$$

We define $\rho_{d-1}(x, y) := ((x_{d-1}, x_d), (y_{d-1}, y_d))$ and let $\gamma^{(d-1)} = (\rho_{d-1})_\#\gamma$, $\gamma^{(d-1)}_K = (\rho_{d-1})_\#\gamma_K$ be the marginals on $((x_{d-1}, x_d), (y_{d-1}, y_d))$. Note that both integrals depend only on x_{d-1} and y_{d-1} which allows us to write the following string of inequalities

$$\int |x_{d-1} - y_{d-1}|^2 \, d\gamma^{(d-1)} \le \int |x_{d-1} - y_{d-1}|^2 \, d\gamma^{(d-1)}_K$$

$$\int \left(\int |x_{d-1} - y_{d-1}|^2 \, d\gamma^{(d-1)}_{(x_d,y_d)} \right) d\gamma^d(x_d, y_d) \le \int \left(\int |x_{d-1} - y_{d-1}|^2 \, d\gamma^{(d-1)}_{(x_d,y_d),K} \right) d\gamma^d_K(x_d, y_d)$$

Recall that $(\pi_d)_\#\gamma = (\pi_d)_\#\gamma_K = \gamma^d$, we can rewrite the previous inequality as

$$\int \left(\int |x_{d-1} - y_{d-1}|^2 \, d\gamma^{d-1}_{(x_d,y_d)}(x_{d-1}, y_{d-1}) \right) d\gamma^d(x_d, y_d) \tag{6}$$

$$\le \int \left(\int |x_{d-1} - y_{d-1}|^2 \, d\gamma^{d-1}_{(x_d,y_d),K}(x_{d-1}, y_{d-1}) \right) d\gamma^d(x_d, y_d) \tag{7}$$

The impulse here is to invoke the optimality and uniqueness of $\gamma^{d-1}_{(x_d,y_d),K}$ as in the previous paragraph to finally claim that $\gamma^{d-1}_{(x_d,y_d)} = \gamma^{d-1}_{(x_d,y_d),K}$. However, to do so we must first verify that $\gamma^{d-1}_{(x_d,y_d)}$

and $\gamma^{d-1}_{(x_d,y_d),K}$ have the same marginals. That is, we must verify that they are couplings with the same source and target measures. Once proved that the two conditionals have similar marginals, we can invoke the optimality of $\gamma^{(d-1)}_{(x_d,y_d),K}$ as it has been constructed to provide the optimal transport from x_{d-1} to y_{d-1} once (x_d, y_d) is fixed. At this juncture inequality (6) poses a contradiction as even though $\gamma^{(d-1)}_{(x_d,y_d),K}$ performs better than $\gamma^{(d-1)}_{(x_d,y_d)}$ for almost all (x_d, y_d), on average it seems like $\gamma^{(d-1)}_{(x_d,y_d)}$ is not worse than $\gamma^{(d-1)}_{(x_d,y_d),K}$. Thus the two results coincide for almost all (x_d, y_d). In light of Lemma 1, for fixed (x_d, y_d) this means showing for all $\phi \in C_c(\mathbb{R})$ that

$$\int \phi(x_{d-1}) d\gamma^{d-1}_{(x_d,y_d)} = \int \phi(x_{d-1}) d\gamma^{d-1}_{(x_d,y_d),K}$$

as well as

$$\int \phi(y_{d-1}) d\gamma^{d-1}_{(x_d,y_d)} = \int \phi(y_{d-1}) d\gamma^{d-1}_{(x_d,y_d),K}.$$

For the time being, we focus on ϕ being a functions of x_{d-1}. Similar proof for y_{d-1} will follow. As it will suffice to prove the equality for γ^d-a.e. (x_d, y_d), using Lemma 2, we can settle with proving for all $\psi \in C_c(\mathbb{R} \times \mathbb{R})$ that

$$\int \psi(x_d, y_d) \left(\int \phi(x_{d-1}) d\gamma^{d-1}_{(x_d,y_d)} \right) d\gamma^d(x_d, y_d) = \int \psi(x_d, y_d) \left(\int \phi(x_{d-1}) d\gamma^{d-1}_{(x_d,y_d),K} \right) d\gamma^d(x_d, y_d).$$

But this equality is equivalent to

$$\int \psi(x_d, y_d)\phi(x_{d-1}) d\gamma^{d-1} = \int \psi(x_d, y_d)\phi(x_{d-1}) d\gamma^{d-1}_K.$$

Recall that proving that $\gamma^{(d-1)}_{(x_d,y_d)}$ and $\gamma^{(d-1)}_{(x_d,y_d),K}$ have same marginals is equivalent to proving the above expression for alll ϕ and ψ. A priori, we are considering an expression of three variables: x_d, y_d, x_{d-1}, and there may be differing correlation structure in γ^{d-1} and γ^{d-1}_K that would cause the above equality to not hold. However, we have previously shown that the projection of γ^{d-1} and γ^{d-1}_K onto (x_d, y_d) is the same measure: γ^d. Moreover, γ^d is given by a transport map $y_d = T_d(x_d)$. Thus, $\psi(x_d, y_d)\phi(x_{d-1})$ is just a function of (x_{d-1}, x_d). We have that γ^{d-1} and γ^{d-1}_K have the same projection onto (x_{d-1}, x_d), namely the marginal of μ on (x_{d-1}, x_d). Thus, the above equality holds, so by the Lemma 2, we have for γ^d-a.e. (x_d, y_d) that $\int \phi(x_{d-1}) d\gamma^{d-1}_{(x_d,y_d)} = \int \phi(x_{d-1}) d\gamma^{d-1}_{(x_d,y_d),K}$ for all $\phi \in C_c(\mathbb{R})$. By applying Lemma 1 we get that the marginals of $\gamma^{d-1}_{(x_d,y_d)}$ and $\gamma^{d-1}_{(x_d,y_d),K}$ on x_{d-1} are equal. We conclude that the marginals of $\gamma^{d-1}_{(x_d,y_d)}$ and $\gamma^{d-1}_{(x_d,y_d),K}$ on y_{d-1} are equal by an analogous argument. The main difference is that we must use the assumption (H-target) to get a Brenier map $x_d = T_d^{-1}(y_d)$, but then the argument proceeds identically.

To summarize, we have shown that $\gamma^{d-1}_{(x_d,y_d)}$ and $\gamma^{d-1}_{(x_d,y_d),K}$ have the same marginals for γ^d-a.e. (x_d, y_d). By the construction of the KR map, we then have for γ^d-a.e. (x_d, y_d) that

$$\int |x_{d-1} - y_{d-1}|^2 \, d\gamma^{d-1}_{(x_d,y_d),K}(x_{d-1}, y_{d-1}) - \int |x_{d-1} - y_{d-1}|^2 \, d\gamma^{d-1}_{(x_d,y_d)}(x_{d-1}, y_{d-1}) \leq 0.$$

But rearranging (6) shows that integrating a nonpositive function results in a nonnegative value. The only way this can occur is if for γ^d-a.e. (x_d, y_d) we have

$$\int |x_{d-1} - y_{d-1}|^2 \, d\gamma^{d-1}_{(x_d,y_d),K}(x_{d-1}, y_{d-1}) = \int |x_{d-1} - y_{d-1}|^2 \, d\gamma^{d-1}_{(x_d,y_d)}(x_{d-1}, y_{d-1}).$$

By the uniqueness of 1-dimensional optimal transport with quadratic cost, we then have for γ^d-a.e. (x_d, y_d) that $\gamma^{d-1}_{(x_d,y_d),K} = \gamma^{d-1}_{(x_d,y_d)}$. In turn, this forces that $\gamma^{d-1} = \gamma^{d-1}_K$.

Claim: If $\gamma^h = \gamma_K^h$ for $1 < h < d$, then $\gamma^{h-1} = \gamma_K^{h-1}$

We now proceed inductively. The essence of this argument is identical to that of the previous claim, just with some additional notational bookkeeping. Using the optimality of $(\pi_k)_\#\gamma_K$ for $k \geq h$ and (4), we can get the following inequalities

$$\begin{aligned}
\sum_{k\geq h} \int \lambda_k(\epsilon)|x_k - y_k|^2 \, d(\pi_k)_\#\gamma_K(x_k, y_k) + \sum_{k<h} \lambda_k(\epsilon) \int (x_k - y_k)^2 d\gamma_\epsilon(x_k, y_k) \\
\leq \sum_{k\geq h} \int |x_k - y_k|^2 \, d(\pi_k)_\#\gamma_\epsilon(x_k, y_k) + \sum_{k<h} \lambda_k(\epsilon) \int (x_k - y_k)^2 d\gamma_\epsilon \\
= \int c_\epsilon d\gamma_\epsilon \leq \int c_\epsilon d\gamma_K \\
= \sum_{k\geq h} \int |x_k - y_k|^2 \, d(\pi_k)_\#\gamma_K(x_k, y_k) + \sum_{k<h} \lambda_k(\epsilon) \int (x_k - y_k)^2 d\gamma_K.
\end{aligned}$$

As before, by getting rid of the common terms, dividing by $\lambda_{h-1}(\epsilon)$, and passing to the limit

$$\int c^{(h-1)} d\gamma \leq \int c^{(h-1)} d\gamma_K \, .$$

Again, we disintegrate the marginals γ^{h-1} and γ_K^{h-1} in a sequence of conditionals until we finally have the integral with respect to (x_{h-1}, y_{h-1}) through the measure $\gamma^{h-1}_{(x_{d:h}, y_{d:h})}$ and $\gamma^{h-1}_{(x_{d:h}, y_{d:h}),K}$. The task of proving that these two conditional measures have same marginal distribution along x_{h-1} and y_{h-1} is achieved via the test functions of the form

$$\psi(x_h, \ldots, x_d, y_h, \ldots, y_d)\phi(x_{h-1})$$

and

$$\psi(x_h, \ldots, x_d, y_h, \ldots, y_d)\phi(y_{h-1}) \, .$$

To apply the same trick in the first case, we replace the $(y_h, \ldots, y_d)$ variables by $(T_K(x_h), \ldots, T_K(x_d))$. In the other case, we need to invert the Knothe-Rosenblatt map which is straightforward as each y_k for $k \in \{h, \ldots, d\}$, is a 1-d monotone transformation of x_k conditioned on $x_{k+1:d}$ which have already been calculated. As before, in the end we obtain that $\gamma^{h-1} = \gamma_K^{h-1}$.

Using this recursive routine, we ultimately prove that $\gamma = \gamma_K$.

□

References

Carlier, G., Galichon, A., and Santambrogio, F. (2010). From knothe's transport to brenier's map and a continuation method for optimal transport. *SIAM Journal on Mathematical Analysis*, 41(6):2554–2576.

Knothe, H. (1957). Contributions to the theory of convex bodies. *Michigan Mathematical Journal*, 4(1):39–52.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The annals of mathematical statistics*, 23(3):470–472.

Villani, C. (2003). *Topics in optimal transportation*, volume 58. American Mathematical Soc.

Villani, C. (2009). *Optimal transport: old and new*, volume 338. Springer.